

# Synthesis and processing of tRNA-related SINE transcripts in *Arabidopsis thaliana*

Thierry Pélissier, Cécile Bousquet-Antonelli, Laurence Lavie and Jean-Marc Deragon\*

CNRS UMR 6547 BIOMOVE and GDR 2157, Université Blaise Pascal Clermont-Ferrand II,  
63177 Aubière Cedex, France

Received May 14, 2004; Revised and Accepted July 15, 2004

## ABSTRACT

Despite the ubiquitous distribution of tRNA-related short interspersed elements (SINEs) in eukaryotic species, very little is known about the synthesis and processing of their RNAs. In this work, we have characterized in detail the different RNA populations resulting from the expression of a tRNA-related SINE S1 founder copy in *Arabidopsis thaliana*. The main population is composed of poly(A)-ending (pa) SINE RNAs, while two minor populations correspond to full-length (fl) or poly(A) minus [small cytoplasmic (sc)] SINE RNAs. Part of the poly(A) minus RNAs is modified by 3'-terminal addition of C or CA nucleotides. All three RNA populations accumulate in the cytoplasm. Using a mutagenesis approach, we show that the poly(A) region and the 3' end unique region, present at the founder locus, are both important for the maturation and the steady-state accumulation of the different S1 RNA populations. The observation that primary SINE transcripts can be post-transcriptionally processed *in vivo* into a poly(A)-ending species introduces the possibility that this paRNA is used as a retroposition intermediate.

## INTRODUCTION

Short interspersed elements (SINEs) are non-autonomous retroposons that have successfully spread within the genome of almost all eukaryotes [reviewed in (1–3)]. SINEs are ancestrally related to functionally important RNAs, such as tRNA, 5S rRNA and 7SL RNA. They possess an internal promoter that can be recognized by the RNA polymerase III (polIII) enzyme complex, and are usually organized in a monomeric or dimeric structure. Monomeric tRNA-related SINE families are present in the genomes of species from all major eukaryotic lineages and this organization is, by far, the most frequent. These elements are composed of a 5' tRNA-related region and a central region of unknown origin, followed by a stretch of homopolymeric adenosine residues or other simple repeats (2,4,5). Although the length of SINE A-rich regions can vary

from a few nucleotides [oligo(A)] up to more than 50 residues, the term poly(A) is commonly used to design this region. The rodent ID (6) and lemur CYN-SINE (also called t-SINE) (7,8) families are exceptions to this organization since these SINEs lack the tRNA-unrelated region and are composed exclusively of the tRNA-related portion followed by simple repeats. tRNA-related SINEs can also be dimeric such as Twin SINEs from *Culex pipiens* (9) and RathE2 SINEs from *Arabidopsis thaliana* (10). The SINE3 family from zebrafish is the only characterized family composed of SINEs with a 5S-related region (instead of a tRNA-related region), followed by a central region of unknown origin and 3'-terminal repeats (11). So far, 7SL-derived SINEs have been found only in mammalian species. They are composed of a 7SL-derived region followed by a poly(A) tail and can be either monomeric (B1 family) or dimeric (Alu family) (12,13). In mammals, the two monomers of dimeric SINEs can also come from two different ancestral sequences. For example, the Alu type II (14) and Tu (15) SINEs have a tRNA-related left monomer and a 7SL-related right monomer while B1-dID SINEs have exactly the opposite organization (16). SINEs likely retropose by target site-primed reverse transcription (TPRT) using the enzymatic machinery of long interspersed elements (LINEs) (17–20). Two different types of the SINE/LINE partnership apparently exist, one based on a shared 3' end motif (20) and the other based on the presence of a common poly(A) tract (19,21).

The biological significance of polIII SINE transcription is an old but still unresolved question. Obviously, SINE transcription is needed to produce retroposition intermediates, and the amount and structure of SINE transcripts will certainly influence the rate of SINE retroposition and, consequently, its mutagenic effect on genomes. Apart from this role, animal 7SL and tRNA-related SINE RNAs have been proposed to work as stress-induced non-coding RNA regulators [reviewed in (22)]. PolIII SINE transcription is induced by several cellular stresses such as heat shock, treatment with cycloheximide, DNA-damaging agents and viral infections (23–33). Under these conditions, SINE RNA can participate in cell recovery by interacting with the PKR kinase to prevent this enzyme from phosphorylating the EIF2- $\alpha$  subunit and blocking translation (22,34) and/or by enabling the cells to selectively stimulate the translation of newly synthesized

\*To whom correspondence should be addressed. Tel: +33 4 73 40 77 52; Fax: +33 4 73 40 77 77; Email: j-marc.deragon@univ-bpclermont.fr

Present address:

Laurence Lavie, Department of Human Genetics, Medical School, University of Saar, 66421 Homburg, Germany

mRNAs in a yet uncharacterized PKR-independent pathway (27). The action of SINE RNAs was apparently conserved in divergent eukaryotic lineages, since an increase in SINE transcription was shown to occur in mice and silkworms subjected to non-lethal stresses (26,31). Therefore, a good understanding of the biology of SINE RNAs (their production, accumulation, maturation and the identification of their proteic partners) is needed to apprehend their mutagenic effects as retroposition intermediates and their potential regulatory effects as non-coding RNA regulators.

The production, maturation and accumulation of SINE RNAs have been studied in detail only for the mammalian 7SL-related SINE Alu and B1 [reviewed in (35)]. Following polIII transcription, two distinct species of Alu and B1 RNAs can be observed; a full-length (fl) RNA and a smaller species called small cytoplasmic (sc) RNA. Since 7SL-related SINEs do not possess an internal terminator for polIII-directed transcription, the flRNA includes, in addition to the SINE sequence and its poly(A), a portion of the flanking 3' genomic insertion site. The half-life of Alu and B1 flRNAs is relatively low (~0.5 h for flAlu RNAs) (36,37). Part of the flRNA population is exported to the cytoplasm while another part is processed in the nucleus to form stable scRNAs (half-life of 3–5 h) that are then exported to the cytoplasm (35–43). scRNAs lack the poly(A) region and the sequence corresponding to the 3' genomic insertion site (in the case of Alu, only the left monomer is stabilized in an scRNA species). Since scRNAs are proposed to be inactive in retroposition [due to the lack of the poly(A) region] (19,21), the maturation of fl to scRNAs is potentially very important to limit the availability of the RNA retroposition intermediate (i.e. the flRNA) and to down-regulate 7SL-related SINE retroposition (35). The impact, if any, of the cytoplasmic accumulation of scRNA on cellular function is unknown.

Although tRNA-related SINEs are much more widely distributed in eukaryotes than 7SL-related SINEs, very little is known about the maturation of their RNAs and this knowledge is limited to mammalian SINE families. In mammals, most of the known tRNA-related SINE families (10 out of 14), including members of the B2 superfamily, have an internal polIII terminator signal (TCT<sub>3-6</sub>) preceded by at least one polyadenylation signal (AATAAA) [reviewed in (44)]. Therefore, transcription of these SINEs stops at the end of the element and a poly(A) region is apparently added post-transcriptionally by an unknown mechanism (44–46), a maturation process very different from the one described for 7SL-related SINEs. For all the other mammalian, as well as non-mammalian, tRNA-related SINE families, the poly(A) region is not added post-transcriptionally but encoded by the element (see above). This includes the rodent tRNA-related ID family for which a maturation mechanism similar to that acting on 7SL-related SINE is apparently responsible for the conversion of a large set of primary ID transcripts into poly(A) minus RNA forms (6). The BC1 gene is a 'founder' (i.e. an active) copy responsible for the production of a significant portion of ID repeats, and its transcript is found in high abundance in rodent brain and in significantly lower abundance in the germ line (6,47,48). In contrast to other ID transcripts, BC1 flRNA is highly stable and is not efficiently modified or processed, accounting only for a small fraction of the detected ID poly(A) minus scRNAs (6). Differences in

processing between BC1 and other ID primary RNAs could have contributed to the success of BC1 as a founder locus. Nothing is known about the processing of non-mammalian tRNA-related SINE transcripts. Transcripts from the silkworm SINE Bm1 have been shown to produce, in addition to the full-length or near full-length RNA, smaller RNA species, one of which was called scBm1 (25). However, these smaller RNA species result from the presence in the SINE portion of leaky internal polIII terminators and are not post-transcriptionally processed from the full-length RNA (25).

*A.thaliana* and *Brassica* species contain several different families of tRNA-related SINEs, called RathE1, RathE2, RathE3, S1 and S2 [(10,49,50) and J. M. Deragon, unpublished data]. In this work, we have characterized in detail for the first time the different RNA populations resulting from the maturation of tRNA-related SINE RNAs. We show that a full-length S1 SINE RNA, produced from a founder locus in *Arabidopsis*, is processed into three RNA populations, including a poly(A) minus scSINE RNA, a situation that is surprisingly similar to the one described for mammalian 7SL-related SINEs. Using a mutagenesis approach, we also show the importance of the 3' end and poly(A) regions on the maturation and accumulation of S1 RNAs.

## MATERIALS AND METHODS

### Recombinant DNA plasmids and transformation

The binary plasmid vectors used in this work were constructed as follows. The original pBI121 vector (Clontech) was modified by exchanging the complete region between the T-DNA left and right borders with an EcoRI–HindIII DNA fragment containing a phosphinotrycine acetyltransferase (glufosinate resistance) gene, expressed from the 35S promoter of the cauliflower mosaic virus. For the C3 transgene, an AscI–EcoRI PCR fragment was amplified using, as template, the na7-3 construct [containing the *Brassica napus* Na7 locus (51)]. For the C1C6 and C5C6 transgenes, a KpnI–AscI PCR fragment containing the proximal 110 bp 7SL enhancer sequence was first amplified using, as template, the cloned 'I-MAS' tomato 7SL gene (52). Next, an AscI–EcoRI PCR fragment was amplified using, as template, either the Na7-3 or Na7-8 [that is similar to Na7-3, but bears mutations within the A box of the S1 element (51)] construct. After digestion with the AscI restriction enzyme, the 7SL and Na7-specific fragments were ligated to generate KpnI–EcoRI fragments representing the C1C6 and C5C6 transgenes. After digestion with the suitable enzymes, the AscI–EcoRI (C3 transgene) or KpnI–EcoRI (C1C6, C5C6 transgenes) fragment was introduced into the modified pBI121 transformation vector. To create the U6Na7 transgene, the 90 bp proximal enhancer of the *A.thaliana* U6 gene (accession no. X52528) was amplified as a 5'HindIII–3'-blunt-end fragment and ligated to a 5'-blunt-end–3'EcoRI fragment, amplified by PCR using the Na7-3 construct as template and corresponding to the Na7 SINE element and its 3'-specific region. A HindIII–EcoRI fragment representing the U6Na7 transgene was then introduced into the HindIII–EcoRI-cut pPZP111 transformation vector that bears a kanamycin (nptII) resistance gene as a plant selection marker (53). The mutated Na7 transgenes were constructed by PCR using the U6Na7 transgene

as template and the appropriate primers (available on request) and introduced into the pPZP111 vector (see above) for cell transformation. The identity of all the constructs was verified by sequencing, using an automated capillary sequencer (CEQ2000, Beckman Coulter) and the DTCS sequencing kit (Beckman Coulter).

*Arabidopsis thaliana* Col0 plant or T87 suspension culture transformations (with the pBI121 or the pPZP111 constructs, respectively) were performed as previously described (54,55). Following standard antibiotic selection, several independent transgenic plant and cell lines (at least three) were established for each construct. They were designated using the name of the transgene construct, followed by the number representing the independent primary transformant that was propagated and analyzed. Three week old transformed plants were transferred into the greenhouse to allow seeds formation. Cell lines were maintained under constant kanamycin selection (20 µg/ml) as liquid suspensions and as calli growing on solid media, as previously described (54).

#### Total RNA isolation, northern blot analysis and quantification

Total RNA was extracted as described in (56) using aerial tissues of 10–15 day old plants, growing *in vitro* on standard MS medium (Sigma M-5519) supplemented with 3% (w/v) sucrose, 0.8% (w/v) bacto agar and 15 µg/ml of glufosinate, or using aliquots from cell suspension cultures (typically obtained by filtration of 5 ml of 3–4 day old subculture) or from 15 day old subcultured calli. For northern blot analysis, RNA samples were heat-treated in 1.5 volume of standard formamide buffer and loaded onto a 7% polyacrylamide (PAA) gel (19:1 acrylamide:bis-acrylamide) containing 7.5 M urea and 1× TBE buffer and separated by electrophoresis. The samples were electroblotted to hybond-XL membranes (Amersham-Pharmacia Biotech) and fixed by UV crosslinking. Pre-hybridization and hybridization was carried out in 50% formamide, 5× SSC, 1% SDS, 1× Denhardt solution, 100 µg/ml of sheared, denaturated hering DNA, 10% (w/v) Dextran sulfate at 42°C. Filters were washed with 1× SSC, 1% SDS for 15 min at 42°C and then 15 min at 65°C, followed by two washes with 0.1× SSC, 1% SDS for 15 min at 65°C. For the Na7 probe, the washing steps were done at 60°C instead of 65°C. Signals were visualized by autoradiography or by using a phosphorimager (Molecular Imager FX, Bio-Rad) for quantification. Three independent transformed lines were analyzed for each construct and gave similar results as those presented in Figures 1, 3 and 4.

#### Cytoplasmic fraction separation

To isolate the cytoplasmic fraction, protoplasts were prepared from 50 ml of 3–4 day old suspension cultures. Cells were collected in a sterile falcon tube and, after cell sedimentation and removal of the supernatant, they were resuspended in 40 ml of the digestion mixture (0.7 M mannitol, 5 mM MES–KOH pH 5.6, 50 mg Macerozyme R10, 300 mg cellulase R10, 5 mM CaCl<sub>2</sub>, 0.05% BSA, 2.5 mM 2-mercaptoethanol) and transferred to a falcon plate. Following a 30 min incubation under vacuum (600 mbar) and a 6 h incubation under gentle agitation, the protoplast suspension was filtrated

through a 40 µm sieve. For all subsequent manipulations the protoplasts were kept at 4°C. The suspension was centrifuged for 5 min at 200 g, and protoplasts were resuspended in 20 ml of resuspension buffer (0.5 M sorbitol, 10 mM MES–KOH pH 5.5, 1 mM CaCl<sub>2</sub>). The centrifugation step was repeated once and the protoplast pellet was resuspended in 3 ml of resuspension buffer, and then carefully applied on top of a 3 ml sucrose 22% cushion. Following a 5 min centrifugation at 300 g, intact protoplasts, localized at the sucrose interphase, were collected and diluted with resuspension buffer to a final volume of 5 ml. One-tenth of the resulting suspension was analyzed as total RNA and the remaining volume processed further to isolate the cytoplasmic fraction as follows. The protoplast suspension was adjusted to 0.4% Triton X-100, gently mixed, incubated 5 min on ice, and centrifuged for 10 min at 1500 g. The supernatant was decanted and kept as the cytoplasmic fraction. RNA isolation was done as described above and a 1/9 volume of the purified cytoplasmic RNA (i.e. equivalent to the 0.1 volume of the protoplast suspension used to purify the total RNA fraction) was analyzed.

#### 3' RACE experiments

Of the total RNA 2 µg was dephosphorylated by treatment with 1.5 U of shrimp alkaline phosphatase (Roche) at 37°C for 1 h, followed by a 15 min step at 60°C to inactivate the enzyme. An arbitrary 5' phosphorylated 20 base oligonucleotide (primer B), carrying at the 3' end a blocking group to prevent self-ligation, was then ligated to the dephosphorylated RNA using T4 RNA ligase (Roche). This ligation step was carried out in 1× T4 RNA ligase buffer (Roche) containing 40 U of RNase inhibitor (RNasin, Promega), 14% PEG 8000, 10 µM of the primer B and 20 U of T4 RNA ligase in a final volume of 40 µl at 22°C for 16 h. Following ligation, the RNA sample was purified on a Chroma-Spin TE-100 column (Clontech), previously equilibrated with DEPC-treated water. One quarter of the eluted fraction was then subjected to cDNA synthesis using a 20 base oligonucleotide (primer C) complementary to primer B and the Expand reverse transcriptase, according to the manufacturer's instructions (Roche). The first round of PCR was performed on a 2 µl aliquot of the 20 µl cDNA mixture, using a combination of the primer C and the Na7-specific primer 1 (position +4 to +23 relative to the SINE S1 transcription initiation site), and the Goldstar *Taq* polymerase (Eurogentec) under standard conditions. A second round of PCR was done using 1.5 µl of a 1/20 dilution of the first PCR as template, and the primer C in combination with the nested Na7-specific primer 2 (position +41 to +60), primer 3 (+69 to +82) or primer 4 (+110 to +127). After electrophoresis, PCR products were excised from the gel and sub-cloned into the PGEM-Teasy vector (Promega). Sequencing was performed as described above.

#### Secondary structure analysis

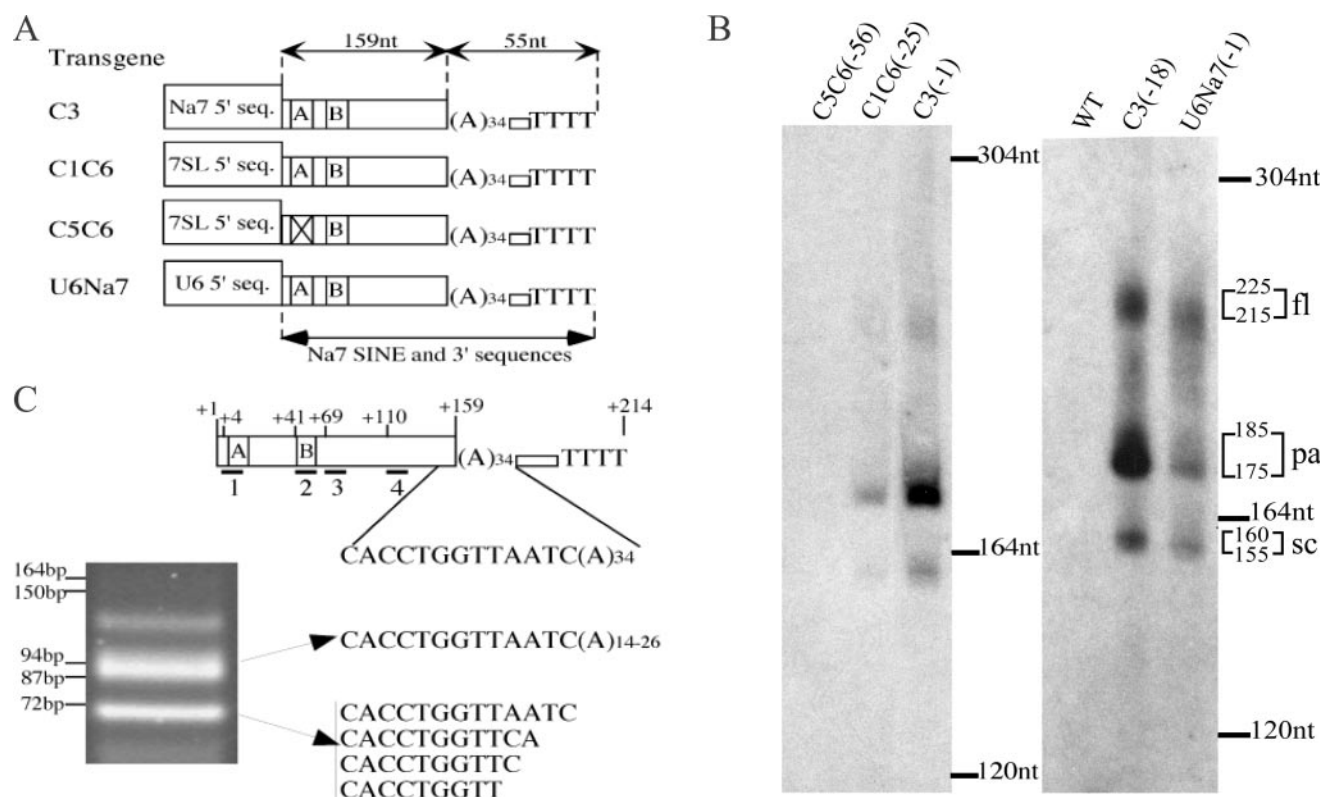
Predicted secondary structures were performed using the RNA mfold version 3.1 (<http://www.bioinfo.rpi.edu/applications/mfold>) (57) and the default thermodynamic parameters (58).

## RESULTS

**tRNA-like SINE elements are post-transcriptionally regulated *in vivo***

The S1 element is a plant SINE that is widely distributed among *Cruciferae* but is absent in *A.thaliana* (50). We introduced an S1 locus from *B.napus*, called Na7, in either *A.thaliana* or cell suspension lines (Figure 1). The SINE sequence at the Na7 locus is identical to the Ea subfamily consensus and is the only S1 copy found to be conserved in all the *Brassica* species tested (59,60). This locus also contains strong transcriptional *cis*-enhancer motifs, upstream of the SINE, that allow this element to partially escape transcriptional repression (51,59,61,62). Therefore, Na7 presents all characteristics of an S1 founder locus. The steady-state accumulation of three S1 RNA populations is detected by northern blot analyses of total RNA extracted from C3 transgenic plant lines, containing the Na7 locus with its natural enhancer

(Figure 1A and B, lanes 3 and 5). The major population contains RNAs of ~175–185 nt, while the two minor populations contain RNAs of ~215–225 and 155–160 nt. To test the impact of the 5' flanking region on the Na7 expression *in vivo*, the proximal enhancer regions from the tomato 7SL or *A.thaliana* U6 gene were substituted for the Na7 5' enhancer and the resulting transgenes (Figure 1A) introduced into the *Arabidopsis* plant and cell lines, respectively. Although the Na7 expression is evidently lower in the C1C6 and U6Na7 transgenic lines than in the C3 transgenic plant lines, the same three RNA populations are detected (Figure 1A and B, lanes 2 and 6). Mutations in the A-box, shown to abolish *in vitro* Na7 polIII transcription (51), prevent the detection of S1 RNA (C5C6 plant lines, Figure 1A and B, lane 1), confirming that the three hybridizing RNA populations result exclusively from S1-driven polIII transcription. Using tobacco nuclear extracts, we observed previously in a time course experiment that the full-length S1 (flS1) RNA produced *in vitro* was transiently



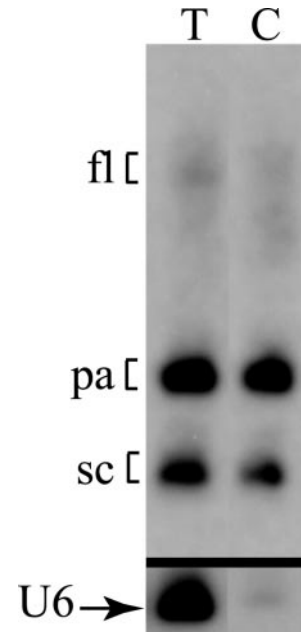
**Figure 1.** Characterization of stable S1 RNA forms originating from a founder locus (Na7). (A) Schematic representation of the SINE transgenes. The C3 transgene contains the *B.napus* Na7 locus including, in addition to the S1 element, its 5' proximal 150 bp region and its 3' specific region, up to the polIII terminator (four T's) (51). Tomato 7SL or *A.thaliana* U6 enhancer sequences were substituted for the Na7 5' proximal region in the C1C6, C5C6 or U6Na7 constructs. The A- and B-boxes of the S1 internal polIII promoter are shown. For the C5C6 transgene, mutations that prevent S1 polIII transcription (51) were introduced within the A-box. These constructs were introduced into *A.thaliana* Col0 plants (for C3, C1C6 and C5C6) or cell (for U6Na7) lines. (B) Northern blot detection of different S1 RNA populations. For each transgene, at least three independent transgenic lines were tested and gave similar results to those reported here. The transgenic lines were identified by the name of the transgene construct, followed by a number within parentheses specific for the independent transformant line that was analyzed. Total RNAs from the C5C6(-56) (lane 1), C1C6(-25) (lane 2), C3(-1) (lane 3) and C3(-18) (lane 5) transgenic plants, the U6Na7(-1) cell suspension (lane 6) and from the wild-type Col0 cell line (WT, lane 4) were fractionated through 7% PAA gels, blotted and hybridized with an  $\alpha$ -<sup>32</sup>P-labeled DNA probe representing the Na7 S1 SINE and its 3' sequences. Size markers were obtained after hybridization of these blots with probes directed against 7SL (304 nt), 5.8S (164 nt) or 5S (120 nt) RNAs. The sizes of the S1-hybridizing bands, deduced from the position of these three markers, are indicated. (C) Molecular characterization of the S1 RNA 3' end. The procedure used to perform the 3' RACE is detailed in the Materials and Methods. The position of the primers and the sequence immediately upstream to the poly(A) region are shown under the map of the Na7 element. An example of a typical PCR result obtained with C3(-1) total RNA [lane 3 in (B)] and the primer 4–primer C couple is shown after separation on 1.8% agarose gel. The two main PCR products represent Na7-specific fragments and the sequences corresponding to the 3' end of these different S1 cDNAs are shown. The faint upper band (>100 bp) contains non-specific products, essentially 5.8S sequences. Independent experiments using C3(-18) and U6Na7(-1) total RNAs as template [lane 5 and 6 in (B)] were performed and gave similar results. For the paS1- and scS1-specific sequences, a total of 27 and 38 clones were analyzed.

converted into a poly(A)-ending intermediate before being processed into a poly(A) minus RNA form (S1). Therefore, according to their northern blot estimated sizes, it appears likely that the largest and the smallest RNA forms observed *in vivo* correspond respectively to flS1 and processed poly(A) minus (scS1) RNAs, while the major S1 RNA population represents a poly(A)-ending form (paS1 RNA).

Using a 3' RACE PCR approach, we were able to amplify, clone and sequence cDNAs corresponding to the three S1 RNA populations detected by northern blot analyses. The major cDNA product corresponds to paS1 RNAs ending with a poly(A) tract of 14 to 26 residues (mean value = 18.5) while the small-size cDNA product corresponds to scS1 RNAs with heterogeneous 3' ends (Figure 1C). The shortest scS1 RNA molecules bear 4 nt less than the longer ones while the two intermediate populations appear to result from the 3' terminal addition of C or CA nucleotides on the shortest product. A few molecules, that resemble the expected flS1 RNA but curiously end by eight to nine A residues, were also detected (data not shown) indicating that a fraction of the S1 primary transcript may have been elongated by a 3' terminal adenylation. The size (larger than the expected 214 nt), and the diffuse pattern of the highest molecular weight signal observed upon northern blot detection of S1 RNAs are both consistent with the existence *in vivo* of flS1 RNAs ending with a variable number of A residues (Figure 1B). Based on these results, we conclude that, after transcription, the S1 precursor transcript is subject to 3' processing, leading to the preferential accumulation of paS1 RNA in addition to the adenylated flS1 and poly(A) minus scS1 forms.

### Cytoplasmic accumulation of the S1 processed products

For 7SL-related Alu and B1 SINEs, the nuclear processing of full-length primary transcripts is correlated with an efficient export of the resulting scRNAs into the cytoplasm, where it accumulates as a stable species (39–41). In our case, previous results using nuclear extracts also suggest that S1 RNAs are processed in the nucleus (51) but whether S1-processed RNAs are exported to the cytoplasm is unknown. To assess the sub-cellular localization of the S1 RNA species, we prepared total and cytoplasmic extracts from U6Na7(-1) cell suspension. Northern blot hybridization of equivalent fractions of our preparations, performed with a Na7-specific probe, reveals that pa and scS1 RNA species mostly, if not exclusively, accumulate in the cytoplasmic compartment (Figure 2). Full-length S1 RNAs are also detected in the cytoplasmic fraction, but since the overall flS1 hybridization signal is weak, the proportion of flS1 RNA present in the cytoplasmic extract is more difficult to evaluate. To test the quality of our cytoplasmic fraction, the filter was re-hybridized with a probe directed against the U6 small nuclear RNA (snRNA), a species known to remain in the nuclear compartment at all times after transcription (63). Figure 2 shows that U6 RNAs are easily detected in the total RNA fraction, but are almost completely absent from the cytoplasmic fraction. A similar result was obtained using a probe directed against the U3 small nucleolar RNA (snoRNA) (data not shown). It is therefore highly unlikely that leakage from the nuclei during cytoplasm preparation is responsible for the identification of S1 RNAs in the cytoplasmic fraction. Moreover, the demonstration that the three

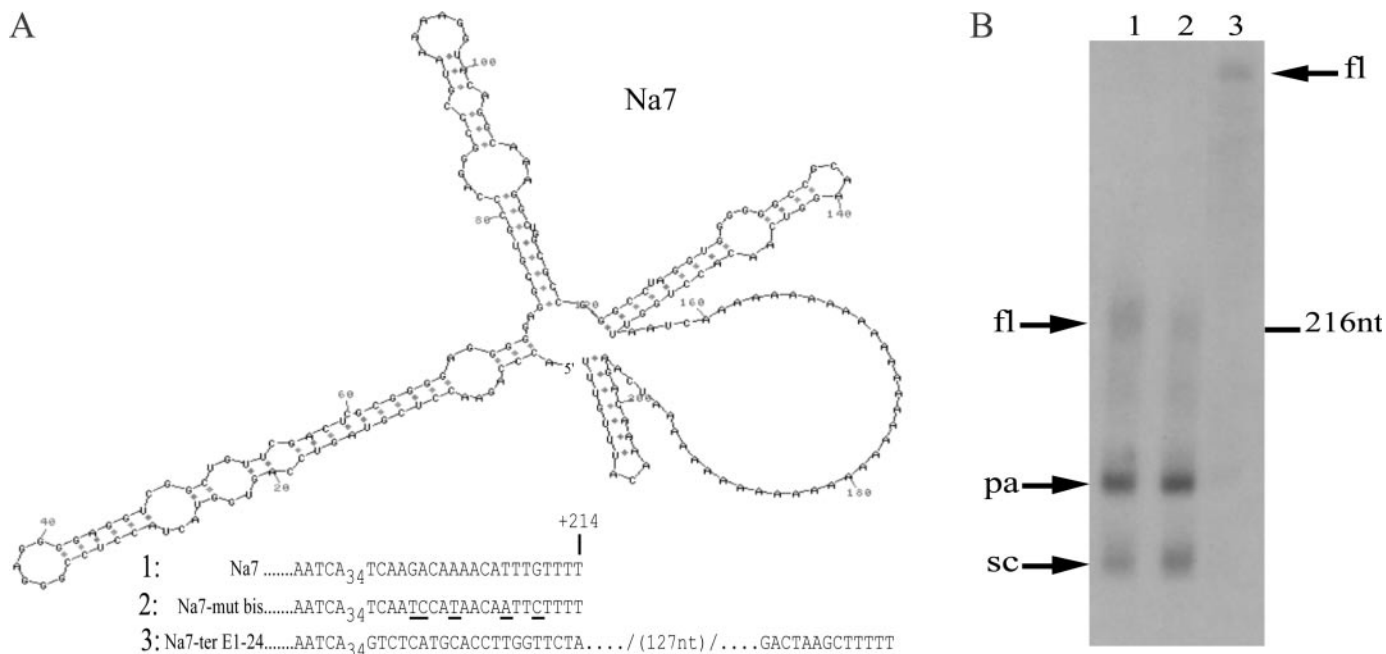


**Figure 2.** Northern blot analysis of cytoplasmic and total RNA fractions. Equivalent loading from the total (lane T) and cytoplasmic (lane C) RNA fractions of the U6Na7(-1) cell line (each representing 0.1 volume of the cell suspension; see Materials and Methods) were analyzed with probes specific for the Na7 SINE RNA and U6 snRNA. The position of the different S1 RNA species and of the U6 RNAs is indicated.

processed S1 RNAs accumulate in the cytoplasm indicates that the paS1 form is not merely a transient nuclear intermediate upstream to the scS1 RNA, but may be associated to a specific function.

### Impact of the unique 3' region at the founder Na7 locus

Only a few SINE loci have the capacity to act as founder loci, (i.e. to be actively amplified). This potentiality requires a combination of several factors affecting SINE transcription as well as the maturation/degradation and retroposition of the resulting SINE transcripts. Sequence and/or structural features of the SINE element can affect this post-transcriptional process, probably by influencing the assembly of a ribonucleoprotein (RNP) complex competent for retroposition (35). However, critical information on a SINE primary transcript also resides outside of the SINE sequence, in the 3' unique portion of the RNA, a region derived from the genomic insertion site. This region was proposed to contribute to SINE RNA stability, processing and reverse transcription (64,65). Computer analysis and preliminary experimental data (data not shown), both indicate that the Na7 SINE RNA sequence can fold mainly into three extended stem-loop structures (Figure 3A). In the full-length transcript, this structure is followed by a single-stranded poly(A) region, and by a short 3' end region that folds into a small terminal stem-loop structure. The presence of a short 3' terminal region that folds into a stem-loop is also characteristic of the mammalian BC1 RNA, the only other known SINE founder locus (66). To evaluate the importance of this terminal structure, we have introduced five mutations within the 3' unique region of



**Figure 3.** Importance of the 3' genomic sequences associated to the SINE. (A) The predicted secondary structure of the Na7 primary transcript, including three U residues of the polIII terminator sequence, is presented. The Na7-mut bis and Na7-ter E1-24 constructs differ from the U6Na7 transgene only by their terminal region downstream of the poly(A) tail. The terminal sequence of these transgenes is reported under the illustration. For the Na7-mut bis transgene, five mutations were introduced into this 3' region (underlined in the sequence) to inhibit the formation of the terminal stem-loop structure, leaving the poly(A) and terminal regions unfolded. For the Na7-ter E1-24 transgene, a longer 3' region specific for a degenerated (non-founder) SINE locus from the *Arabidopsis* RathE1 family (RathE1-24 accession no. AB020746, position 77121–77270) was substituted for the Na7-specific 3' end. (B) Fifteen micrograms of total RNAs extracted from 15 day old calli of the U6Na7(-2) (lane 1), U6Na7-mut bis(-1) (lane 2) and U6Na7-terE1-24(-5) (lane 3) transgenic lines were analyzed with the Na7-specific probe. The position of the full-length- (fl), poly(A)-ending- (pa), small cytoplasmic- (sc) species as well as those of the 216 nt long U3 RNA is shown. For each construct, three independent transformed lines were analyzed and gave similar results.

Na7 (Figure 3A) that are predicted to leave the 3' terminal region unfolded. Northern blot analyses do not reveal a significant modification in the pattern of S1 RNA species accumulation when the mutant construct is compared with the wild-type organization (Figure 3B, compare lanes 1 and 2), indicating that the destabilization of the Na7 3' terminal structure does not strongly alter the processing and steady-state accumulation of the Na7 RNA species. When the Na7 terminator sequence is exchanged for a longer, unrelated region (Na7-ter E1-24 construct, Figure 3A), the pattern of S1 RNA accumulation is significantly modified (Figure 3B, lane 3). The main species detected is the full-length RNA; this form accumulates at a level similar to that observed for the wild-type locus, while the level of the two other RNA species is strongly reduced (Figure 3B). Therefore, the association of a long 3' region to the SINE sequence appears to mainly impede the formation/accumulation of the processed pa and scRNA species.

#### The poly(A) region is an important *cis* factor in S1 RNA maturation

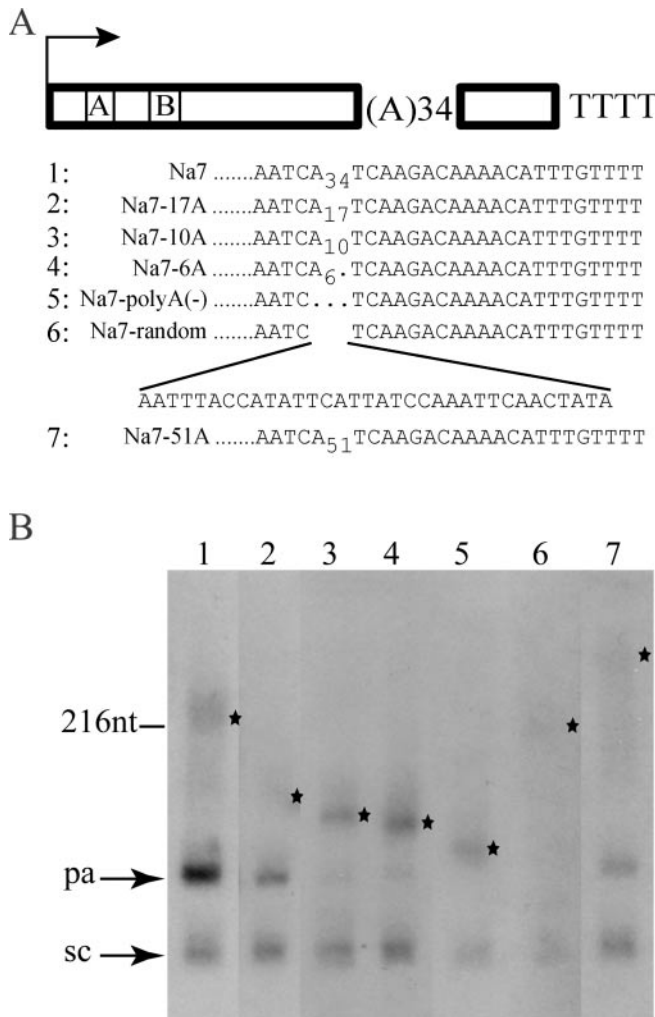
A series of transgenes was constructed to test the importance of the poly(A) region in the SINE RNA maturation process (Figure 4A). First, the length of the poly(A) region was progressively shortened in the different transgenes, from 34 A residues (Na7 wild type) to 17 (Na7-17A), 10 (Na7-10A), 6 (Na7-6A) and up to the complete deletion of the poly(A) region [Na7-poly(A<sup>-</sup>)]. In two other constructs, the size of the

poly(A) region was increased to 51 A residues (Na7-51A) or replaced by a 34 bp heterogeneous sequence that contains only 14 interspersed A residues (Na7-random). For all these modified transgenes, the predicted RNA folding pattern is similar to that of the original Na7 element (Figure 3A and data not shown). The shortening of the poly(A) tail from 34 to 17 A residues does not modify the overall S1 RNA accumulation pattern (Figure 4B, lanes 1 and 2) but further shortening of the poly(A) results in the dramatic reduction of the intermediate RNA species formation and/or accumulation (Figure 4B, lanes 3 to 5). The loss of the intermediate size species is also observed when the poly(A) region is replaced by a 34 base non-poly(A) region (Figure 4B, lane 6). The presence of a poly(A) tail of 51 residues (Figure 4B, lane 7) does not strongly affect the RNA accumulation pattern as compared to the one observed with the original Na7 element. Altogether, these results suggest that a poly(A) tail longer than 10 residues is an important *cis* factor for the formation and/or accumulation of the pa-like species although the shortening, or even the complete deletion of this poly(A) region, does not prevent the accumulation of the smaller scRNA species.

## DISCUSSION

### RNA processing of a tRNA-related SINE

tRNA-related SINEs are ubiquitous in eukaryotes, and represent, by far, the major class of SINEs. We show here that the tRNA-related S1 SINE primary transcript can be processed



**Figure 4.** Impact of the poly(A) tail region. (A) The mutated transgenes are shown under the schematic representation of the Na7 element. (B) Typical results of a northern blot analysis, obtained by the hybridization with the Na7 probe of a blot containing 15 µg of total RNAs extracted from 15 day old calli of the U6Na7(-2) (lane 1), U6Na7-17A(-2) (lane 2), U6Na7-10A(-10) (lane 3), U6Na7-6A(-2) (lane 4), U6Na7-polyA(-)(-4) (lane 5), U6Na7-random(-7) (lane 6) and U6Na7-51A(-5) (lane 7) transgenic lines. The position of the full-length transcript is indicated in each lane, by an asterisk. For each construct, three independent transformed lines were analyzed and gave similar results.

*in vivo* in three different RNA species, a polyadenylated full-length species, a small poly(A) minus species (scS1) and a major product, intermediate in size, corresponding to a poly(A)-ending species (paS1). These results first imply that the capacity to process SINE primary transcripts is conserved between plants and mammals, and between 7SL-related and certain tRNA-related SINE families, although the nature of the processed products can vary. We were also able to observe that small scS1 RNAs present heterogeneous 3' ends; the shortest scS1 RNA molecules bear 4 nt less than the longer ones while the two intermediate populations appear to result from the 3' terminal addition of C or CA nucleotides on the shortest product. The post-transcriptional addition of C or CA, observed here for the first time on a SINE RNA, has been reported for various non-tRNA plant and animal transcripts

and is proposed to be the result of the activity of a nucleotidyltransferase similar to the CCA:tRNA-adding enzyme (67,68). The biological importance of this editing process is unknown, but it could stabilize scS1 RNA as this is probably the case for the mammalian U2 snRNA (67). The presence of several A's at the 3' end of fl-S1 RNA is also intriguing and could result from the action of the SRP 7SL adenylating enzyme (69,70), the poly(A) polymerase enzyme (71) or another adenylation system. But perhaps the most surprising result is the observation that the major processed RNA species is composed of molecules ending with 14 to 26 A residues. These molecules are not merely nuclear transient intermediates upstream to the scS1 formation since they can accumulate in the cytoplasm with the two other RNA species (Figure 2). Also, the formation of sc and paS1 RNAs is not an artifact linked to the expression of a *B.napus* SINE in *A.thaliana*, since preliminary results show that the primary transcript of a putative *Arabidopsis* SINE founder sequence [from the RathE1 family (49)] is also processed in sc and paRNA species when expressed, as a transgene, in the same *Arabidopsis* ecotype (data not shown).

Based on our current knowledge of the SINE/LINE retroposition process, it is tempting to propose a role for the paS1 RNA as a possible retroposition intermediate. In current models, SINE TPRT is proposed to initiate directly on the internal A-tract of flSINE RNA (18,21) or to proceed by template switching from LINE to flSINE RNAs during reverse transcription of the poly(A) tail (72,73). LINEs from the L1 clade [like Bali1, the putative partner of the SINE S1 (74)] have a low ability to retropose non-LINE RNAs (75) and recent *in vitro* studies have shown that L1 reverse transcription is more efficiently initiated at a terminal rather than at an internal A-tract since the presence of a 3' trailer region interferes with the overall efficiency of the reverse transcription step (72). The formation of SINE RNAs ending with a long terminal poly(A) tail could therefore be critical to efficiently compete with the poly(A)-ending LINE RNA and to better access the LINE retroposition machinery. Interestingly, most of the characterized S1 genomic copies have a poly(A) tract of 16 to 24 residues (mean value = 19) (76,77), a situation compatible with a TPRT process initiating at the 3' end of paS1 RNA that mainly displays A-tails of 14 to 26 residues (mean value = 18.5). Although flS1 RNA can also end with a short A-tract, reverse transcription is apparently not initiated at the 3' end of flS1 transcripts since this would lead to the transduction of flanking genomic sequences, a situation observed for LINEs (78) but never for S1 or any other SINEs. Founder loci have the rare potentiality to generate a SINE family. We propose that the capacity to produce a significant amount of paS1 RNA could be an important feature of founder loci and that the nature of the poly(A) region and of the 3' end region is involved in this maturation step. In our case, although the introduced SINE Na7 transgenes were not targeted by silencing modifications like DNA methylation (62,79) and appeared to be stably maintained and expressed for at least four generations after their introduction into the *Arabidopsis* plants, we could not detect evidence for Na7 retroposition (data not shown). This expected result is likely due to the absence of activity of the potential LINE partners present in the *Arabidopsis* genome (80).

### Cis factors involved in SINE RNA maturation

In a first effort to determine the *cis* factors that are important for the paS1/scS1 formation process, mutations were introduced in the 3' terminal stem-loop of the flS1 RNA (Figure 3). Since a similar stem-loop can be found at the flBC1RNA 3' end (66), the only other known SINE founder RNA, it was important to test the impact of this structure on flS1 RNA stability/maturation. Mutations destabilizing this terminal structure did not alter significantly the S1 RNA accumulation pattern, suggesting that a terminal stem-loop is not an essential feature ensuring the correct processing of the primary transcript in pa and/or scRNA species. However, the substitution of the very short Na7 3' end region with a much longer one (from a degenerated *Arabidopsis* SINE locus called RathE1-24) did have a strong negative impact on both pa and scRNA formation, suggesting that the size of the 3' terminal region, more than the structure itself, could be important for proper processing. The fact that similar levels of flS1 are found when using transgenes bearing the short (Na7) or long (Na7-ter E1-24) 3' end regions suggests that flS1 RNAs are affected by this substitution since, by impeding pa and scRNAs formation, we should have observed a higher accumulation of the flS1 species. The fact that we do not see this could be the result of a lower production or an incorrect stabilization and/or processing of the flS1 transcript when using the Na7-ter E1-24 construct.

Another potentially important *cis*-acting factor is the poly(A) region. The poly(A) region was shown recently to be essential for Alu retroposition (19) and poly(A)-binding proteins (PABPs) are known to bind to this region and to participate in the formation of mammalian SINE RNP structures (81,82). The shortening of the Na7 poly(A) region from 34 to 17 A residues does not have a dramatic impact on S1 RNA processing. Further shortening of the poly(A) region drastically reduces the accumulation of an intermediate size species. Increasing the size of the poly(A) region does not alter significantly the S1 maturation process, but the replacement of the poly(A) region with an unrelated 34 bp sequence also prevents the accumulation of an intermediate size species. However, in contrast to the situation of B1 RNA (40), the production of scRNA is not strongly affected by the reduction or even the complete absence of the poly(A) region. These results strongly suggest that a *trans* factor, recognizing a poly(A) of more than 10 bases (possibly a PABP), is needed to stabilize the intermediate (paS1) RNA species but is not essential for the processing of flS1 into scS1 RNA. It is also likely that a set of specific proteins (yet to be characterized) would interact with the internal SINE S1 sequence to form a core RNP complex. For 7SL-related elements, the heterodimeric SRP protein known as SRP 9/14 has been shown to participate in the assembly of a RNP structure (83). The recruitment of these two proteins is dependent on a specific folding of the primary transcripts of 7SL-related SINEs and is probably essential for the stability, processing and/or the retroposition ability of the SINE molecules.

In conclusion, an 'active' S1 founder locus would need to (i) escape transcriptional repression, mainly by using polIII 5' external enhancers (51,62), (ii) have a correct SINE sequence in order to assemble a protective SINE RNP structure and (iii) possess a long poly(A) region and a short 3' end region to favor

the production and cytoplasmic export of the poly(A)-ending retroposition intermediate. A direct SINE/LINE retroposition assay is now being developed in *Arabidopsis* to test this model.

### ACKNOWLEDGEMENTS

We thank Nicolas Gilbert for critical reading of the manuscript. This work was supported by the CNRS (UMR 6547 BIOMOVE and GDR2157), by Université Blaise Pascal and by the European Commission as part of the RIBOREG program (503022).

### REFERENCES

- Deininger, P.L., Moran, J.V., Batzer, M.A. and Kazazian, H.H., Jr (2003) Mobile elements and mammalian genome evolution. *Curr. Opin. Genet. Dev.*, **13**, 651–658.
- Okada, N. (1991) SINEs. *Curr. Opin. Genet. Dev.*, **1**, 498–504.
- Smit, A.F. (1999) Interspersed repeats and other mementos of transposable elements in mammalian genomes. *Curr. Opin. Genet. Dev.*, **9**, 657–663.
- Okada, N. and Ohshima, K. (1995) Evolution of tRNA-derived SINEs. In Marais, R. (ed.), *The Impact of Short Interspersed Elements (SINEs) on the Host Genome*. RG Landes Company, Springer, TX, pp. 61–80.
- Okada, N., Hamada, M., Ogiwara, I. and Ohshima, K. (1997) SINEs and LINEs share common 3' sequences: a review. *Gene*, **205**, 229–243.
- Kim, J., Kass, D.H. and Deininger, P.L. (1995) Transcription and processing of the rodent ID repeat family in germline and somatic cells. *Nucleic Acids Res.*, **23**, 2245–2251.
- Piskurek, O., Nikaido, M., Boeadi, B., Baba, M. and Okada, N. (2003) Unique mammalian tRNA-derived repetitive elements in dermopterans: the t-SINE family and its retrotransposition through multiple sources. *Mol. Biol. Evol.*, **20**, 1659–1668.
- Schmitz, J. and Zischler, H. (2003) A novel family of tRNA-derived SINEs in the colugo and two new retrotransposable markers separating dermopterans from primates. *Mol. Phylogenet. Evol.*, **28**, 341–349.
- Feschotte, C., Fourier, N., Desmons, I. and Mouches, C. (2001) Birth of a retroposon: the Twin SINE family from the vector mosquito *Culex pipiens* may have originated from a dimeric tRNA precursor. *Mol. Biol. Evol.*, **18**, 74–84.
- Lenoir, A., Pélissier, T., Bousquet-Antonelli, C. and Deragon, J.M. (2004) Comparative evolution history of SINEs in *Arabidopsis thaliana* and *Brassica oleracea*: evidence for a high rate of SINE loss. *Cytogenet. Cell Genet.*, In press.
- Kapitonov, V.V. and Jurka, J. (2003) A novel class of SINE elements derived from 5S rRNA. *Mol. Biol. Evol.*, **20**, 694–702.
- Labuda, D., Sinnett, D., Richer, C., Deragon, J.M. and Striker, G. (1991) Evolution of mouse B1 repeats: 7SL RNA folding pattern conserved. *J. Mol. Evol.*, **32**, 405–414.
- Sinnett, D., Richer, C., Deragon, J.M. and Labuda, D. (1991) Alu RNA secondary structure consists of two independent 7 SL RNA-like folding units. *J. Biol. Chem.*, **266**, 8675–8678.
- Daniels, G.R. and Deininger, P.L. (1985) Repeat sequence families derived from mammalian tRNA genes. *Nature*, **317**, 819–822.
- Nishihara, H., Terai, Y. and Okada, N. (2002) Characterization of novel Alu- and tRNA-related SINEs from the tree shrew and evolutionary implications of their origins. *Mol. Biol. Evol.*, **19**, 1964–1972.
- Vasetskii, N.S., Gogolevskaya, I.K., Borodulina, O.R. and Kramerov, D. (1999) A new rodent short retroposon B1-dID. *Mol. Biol.*, **33**, 520–527.
- Weiner, A.M. (2002) SINEs and LINEs: the art of biting the hand that feeds you. *Curr. Opin. Cell Biol.*, **14**, 343–350.
- Weiner, A.M. (2000) Do all SINEs lead to LINEs? *Nature Genet.*, **24**, 332–333.
- Dewannieux, M., Esnault, C. and Heidmann, T. (2003) LINE-mediated retrotransposition of marked Alu sequences. *Nature Genet.*, **35**, 41–48.
- Kajikawa, M. and Okada, N. (2002) LINEs mobilize SINEs in the eel through a shared 3' sequence. *Cell*, **111**, 433–444.
- Boeke, J.D. (1997) LINEs and Alus—the polyA connection. *Nature Genet.*, **16**, 6–7.



22. Schmid, C.W. (1998) Does SINE evolution preclude Alu function? *Nucleic Acids Res.*, **26**, 4541–4550.
23. Liu, W.M., Chu, W.M., Choudary, P.V. and Schmid, C.W. (1995) Cell stress and translational inhibitors transiently increase the abundance of mammalian SINE transcripts. *Nucleic Acids Res.*, **23**, 1758–1765.
24. Kim, C., Rubin, C.M. and Schmid, C.W. (2001) Genome-wide chromatin remodeling modulates the Alu heat shock response. *Gene*, **276**, 127–133.
25. Kimura, R.H., Choudary, P.V. and Schmid, C.W. (1999) Silk worm Bm1 SINE RNA increases following cellular insults. *Nucleic Acids Res.*, **27**, 3380–3387.
26. Kimura, R.H., Choudary, P.V., Stone, K.K. and Schmid, C.W. (2001) Stress induction of Bm1 RNA in silkworm larvae: SINEs, an unusual class of stress genes. *Cell Stress Chaperones*, **6**, 263–272.
27. Rubin, C.M., Kimura, R.H. and Schmid, C.W. (2002) Selective stimulation of translational expression by Alu RNA. *Nucleic Acids Res.*, **30**, 3253–3261.
28. Rudin, C.M. and Thompson, C.B. (2001) Transcriptional activation of short interspersed elements by DNA-damaging agents. *Genes Chromosomes Cancer*, **30**, 64–71.
29. Li, T.H., Kim, C., Rubin, C.M. and Schmid, C.W. (2000) K562 cells implicate increased chromatin accessibility in Alu transcriptional activation. *Nucleic Acids Res.*, **28**, 3031–3039.
30. Li, T.H. and Schmid, C.W. (2001) Differential stress induction of individual Alu loci: implications for transcription and retrotransposition. *Gene*, **276**, 135–141.
31. Li, T., Spearow, J., Rubin, C.M. and Schmid, C.W. (1999) Physiological stresses increase mouse short interspersed element (SINE) RNA expression *in vivo*. *Gene*, **239**, 367–372.
32. Panning, B. and Smiley, J.R. (1994) Activation of RNA polymerase III transcription by adenovirus type 5 and herpes simplex virus type 1. In Maraia, R. (ed.), *The Impact of Short Interspersed Elements (SINEs) on the Host Genome*. R.G. Landes Company, TX, pp. 143–161.
33. Wick, N., Luedemann, S., Vietor, I., Cotten, M., Wildpaner, M., Schneider, G., Eisenhaber, F. and Huber, L.A. (2003) Induction of short interspersed nuclear repeat-containing transcripts in epithelial cells upon infection with a chicken adenovirus. *J. Mol. Biol.*, **328**, 779–790.
34. Gale, M., Jr and Katze, M.G. (1998) Molecular mechanisms of interferon resistance mediated by viral-directed inhibition of PKR, the interferon-induced protein kinase. *Pharmacol. Ther.*, **78**, 29–46.
35. Maraia, R. and Sarrowa, J. (1995) Alu-family SINE RNA: interacting proteins and pathways of expression. In Maraia, R. (ed.), *The Impact of Short Interspersed Elements (SINEs) on the Host Genome*. R.G. Landes company, TX, pp. 163–196.
36. Chu, W.M., Liu, W.M. and Schmid, C.W. (1995) RNA polymerase III promoter and terminator elements affect Alu RNA expression. *Nucleic Acids Res.*, **23**, 1750–1757.
37. Li, T.H. and Schmid, C.W. (2004) Alu's dimeric consensus sequence destabilizes its transcripts. *Gene*, **324**, 191–200.
38. Liu, W.M., Maraia, R.J., Rubin, C.M. and Schmid, C.W. (1994) Alu transcripts: cytoplasmic localisation and regulation by DNA methylation. *Nucleic Acids Res.*, **22**, 1087–1095.
39. Adeniyi-Jones, S. and Zasloff, M. (1985) Transcription, processing and nuclear transport of a B1 Alu RNA species complementary to an intron of the murine alpha-fetoprotein gene. *Nature*, **317**, 81–84.
40. Maraia, R.J. (1991) The subset of mouse B1 (Alu-equivalent) sequences expressed as small processed cytoplasmic transcripts. *Nucleic Acids Res.*, **19**, 5695–5702.
41. Maraia, R.J., Driscoll, C.T., Bilyeu, T., Hsu, K. and Darlington, G.J. (1993) Multiple dispersed loci produce small cytoplasmic Alu RNA. *Mol. Cell. Biol.*, **13**, 4233–4241.
42. Sinnett, D., Richer, C., Deragon, J.M. and Labuda, D. (1992) Alu RNA transcripts in human embryonal carcinoma cells. Model of post-transcriptional selection of master sequences. *J. Mol. Biol.*, **226**, 689–706.
43. Shaikh, T.H., Roy, A.M., Kim, J., Batzer, M.A. and Deininger, P.L. (1997) cDNAs derived from primary and small cytoplasmic Alu (scAlu) transcripts. *J. Mol. Biol.*, **271**, 222–234.
44. Borodulina, O.R. and Kramerov, D.A. (2001) Short interspersed elements (SINEs) from insectivores. Two classes of mammalian SINEs distinguished by A-rich tail structure. *Mamm. Genome*, **12**, 779–786.
45. Kramerov, D.A., Tillib, S.V., Shumyatsky, G.P. and Georgiev, G.P. (1990) The most abundant nascent poly(A<sup>+</sup>) RNAs are transcribed by RNA polymerase III in murine tumor cells. *Nucleic Acids Res.*, **18**, 4499–4506.
46. Bladon, T.S., Fregeau, C.J. and McBurney, M.W. (1990) Synthesis and processing of small B2 transcripts in mouse embryonal carcinoma cells. *Mol. Cell. Biol.*, **10**, 4058–4067.
47. Kim, J., Martignetti, J.A., Shen, M.R., Brosius, J. and Deininger, P. (1994) Rodent BC1 RNA gene as a master gene for ID element amplification. *Proc. Natl Acad. Sci. USA*, **91**, 3607–3611.
48. DeChiara, T.M. and Brosius, J. (1987) Neural BC1 RNA: cDNA clones reveal nonrepetitive sequence content. *Proc. Natl Acad. Sci. USA*, **84**, 2624–2628.
49. Lenoir, A., Lavie, L., Prieto, J.L., Goubely, C., Cote, J.C., Pelissier, T. and Deragon, J.M. (2001) The evolutionary origin and genomic organization of SINEs in *Arabidopsis thaliana*. *Mol. Biol. Evol.*, **18**, 2315–2322.
50. Lenoir, A., Cournoyer, B., Warwick, S., Picard, G. and Deragon, J.M. (1997) Evolution of SINE S1 retrotransposons in Cruciferae plant species. *Mol. Biol. Evol.*, **14**, 934–941.
51. Arnaud, P., Yukawa, Y., Lavie, L., Pelissier, T., Sugiura, M. and Deragon, J.M. (2001) Analysis of the SINE S1 Pol III promoter from Brassica; impact of methylation and influence of external sequences. *Plant J.*, **26**, 295–305.
52. Riedel, L., Volger, U., Luckinger, R., Putz, A., Sanger, H.L. and Wassenegger, M. (1996) Molecular analysis of the gene family of the signal recognition particle (SRP) RNA of tomato. *Plant Mol. Biol.*, **31**, 113–125.
53. Hajdukiewicz, P., Svab, Z. and Maliga, P. (1994) The small, versatile pPZP family of Agrobacterium binary vectors for plant transformation. *Plant Mol. Biol.*, **25**, 989–994.
54. Gallego, M.E., Sirand-Pugnet, P. and White, C.I. (1999) Positive-negative selection and T-DNA stability in Arabidopsis transformation. *Plant Mol. Biol.*, **39**, 83–93.
55. Clough, S.J. and Bent, A.F. (1998) Floral dip: a simplified method for Agrobacterium-mediated transformation of *Arabidopsis thaliana*. *Plant J.*, **16**, 735–743.
56. Logemann, J., Schell, J. and Willmitzer, L. (1987) Improved method for the isolation of RNA from plant tissues. *Anal. Biochem.*, **163**, 16–20.
57. Zuker, M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.*, **31**, 3406–3415.
58. Mathews, D.H., Sabina, J., Zuker, M. and Turner, D.H. (1999) Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J. Mol. Biol.*, **288**, 911–940.
59. Deragon, J.M., Gilbert, N., Rouquet, L., Lenoir, A., Arnaud, P. and Picard, G. (1996) A transcriptional analysis of the S1Bn (*Brassica napus*) family of SINE retrotransposons. *Plant Mol. Biol.*, **32**, 869–878.
60. Tatout, C., Warwick, S.I., Lenoir, A. and Deragon, J.M. (1999) SINE insertion as clade markers for wild Cruciferae species. *Mol. Biol. Evol.*, **16**, 1614–1621.
61. Goubely, C., Arnaud, P., Tatout, C., Heslop-Harrison, J.S. and Deragon, J.M. (1999) S1 SINE retrotransposons are methylated at symmetrical and non-symmetrical positions in *Brassica napus*: identification of a preferred target site for asymmetrical methylation. *Plant Mol. Biol.*, **39**, 243–255.
62. Arnaud, P., Goubely, C., Pelissier, T. and Deragon, J.M. (2000) SINE retrotransposons can be used *in vivo* as nucleation centers for *de novo* methylation. *Mol. Cell. Biol.*, **20**, 3434–3441.
63. Boelens, W.C., Palacios, I. and Mattaj, J.W. (1995) Nuclear retention of RNA as a mechanism for localization. *RNA*, **1**, 273–283.
64. Aleman, C., Roy-Engel, A.M., Shaikh, T.H. and Deininger, P.L. (2000) Cis-acting influences on Alu RNA levels. *Nucleic Acids Res.*, **28**, 4755–4761.
65. Shen, M.R., Brosius, J. and Deininger, P.L. (1997) BC1 RNA, the transcript from a master gene for ID element amplification, is able to prime its own reverse transcription. *Nucleic Acids Res.*, **25**, 1641–1648.
66. Rozhdetsvensky, T.S., Kopylov, A.M., Brosius, J. and Huttenhofer, A. (2001) Neuronal BC1 RNA structure: evolutionary conversion of a tRNA(Ala) domain into an extended stem-loop structure. *RNA*, **7**, 722–730.
67. Cho, H.D., Tomita, K., Suzuki, T. and Weiner, A.M. (2002) U2 small nuclear RNA is a substrate for the CCA-adding enzyme (tRNA nucleotidyltransferase). *J. Biol. Chem.*, **277**, 3447–3455.
68. Williams, M.A., Johzuka, Y. and Mulligan, R.M. (2000) Addition of non-genomically encoded nucleotides to the 3'-terminus of maize mitochondrial mRNAs: truncated rps12 mRNAs frequently terminate with CCA. *Nucleic Acids Res.*, **28**, 4444–4451.

69. Perumal,K., Sinha,K., Henning,D. and Reddy,R. (2001) Purification, characterization, and cloning of the cDNA of human signal recognition particle RNA 3'-adenylating enzyme. *J. Biol. Chem.*, **276**, 21791–21796.
70. Sinha,K.M., Gu,J., Chen,Y. and Reddy,R. (1998) Adenylation of small RNAs in human cells. Development of a cell-free system for accurate adenylation on the 3'-end of human signal recognition particle RNA. *J. Biol. Chem.*, **273**, 6853–6859.
71. Zhao,J., Hyman,L. and Moore,C. (1999) Formation of mRNA 3' ends in eukaryotes: mechanism, regulation, and interrelationships with other steps in mRNA synthesis. *Microbiol. Mol. Biol. Rev.*, **63**, 405–445.
72. Cost,G.J., Feng,Q., Jacquier,A. and Boeke,J.D. (2002) Human L1 element target-primed reverse transcription *in vitro*. *EMBO J.*, **21**, 5899–5910.
73. Brosius,J. (1999) RNAs from all categories generate retrosequences that may be exapted as novel genes or regulatory elements. *Gene*, **238**, 115–134.
74. Lavie,L., Prieto,J.L., Tatout,C., Deragon,J.M. (2002) Bali1A, a possible LINE partner for the SINE S1 in *Brassica napus*. *Genome Lett.*, **3**, 123–130.
75. Wei,W., Gilbert,N., Ooi,S.L., Lawler,J.F., Ostertag,E.M., Kazazian,H.H., Boeke,J.D. and Moran,J.V. (2001) Human L1 retrotransposition: *cis* preference versus *trans* complementation. *Mol. Cell. Biol.*, **21**, 1429–1439.
76. Gilbert,N., Arnaud,P., Lenoir,A., Warwick,S.I., Picard,G. and Deragon,J.M. (1997) Plant S1 SINEs as a model to study retroposition. *Genetica*, **100**, 155–160.
77. Tatout,C., Lavie,L. and Deragon,J.M. (1998) Similar target site selection occurs in integration of plant and mammalian retroposons. *J. Mol. Evol.*, **47**, 463–470.
78. Pickeral,O.K., Makalowski,W., Boguski,M.S. and Boeke,J.D. (2000) Frequent human genomic DNA transduction driven by LINE-1 retrotransposition. *Genome Res.*, **10**, 411–415.
79. Zilberman,D., Cao,X. and Jacobsen,S.E. (2003) ARGONAUTE4 control of locus-specific siRNA accumulation and DNA and histone methylation. *Science*, **299**, 716–719.
80. Ohta,Y., Noma,K., Tsuchimoto,S., Ohtsubo,E. and Ohtsubo,H. (2002) Expression of *Arabidopsis* LINEs from two promoters. *Plant J.*, **32**, 809–818.
81. West,N., Roy-Engel,A.M., Imataka,H., Sonenberg,N. and Deininger,P. (2002) Shared protein components of SINE RNPs. *J. Mol. Biol.*, **321**, 423–432.
82. Muddashetty,R., Khanam,T., Kondrashov,A., Bundman,M., Iacoangeli,A., Kremerskothen,J., Duning,K., Barnekow,A., Huttenhofer,A., Tiedge,H. *et al.* (2002) Poly(A)-binding protein is associated with neuronal BC1 and BC200 ribonucleoprotein particles. *J. Mol. Biol.*, **321**, 433–445.
83. Sarrowa,J., Chang,D.Y. and Maraia,R.J. (1997) The decline in human Alu retroposition was accompanied by an asymmetric decrease in SRP9/14 binding to dimeric Alu RNA and increased expression of small cytoplasmic Alu RNA. *Mol. Cell. Biol.*, **17**, 1144–1151.